

Exact Summatory Functions for Prime k -tuples

J. LaChapelle

Abstract

Exact summatory functions that count the number of prime k -tuples up to some cut-off integer are presented. Related k -tuple analogs of the first and second Chebyshev functions are then defined.

1 Introduction

Interest in counting prime k -tuples for $k \geq 2$ is a familiar story. A notable contribution to the storyline came from Hardy and Littlewood in their influential paper [1] where they conjectured remarkably accurate asymptotics. More contemporary works (see e.g. [2],[3],[4] and references therein) mainly utilize sieve methods to progress, and Yitang Zhang's recent breakthrough [5] has generated considerable excitement.

In this letter we present exact summatory functions that count the number of prime k -tuples up to some cut-off integer. The construction is based on the rather trivial observation that the arithmetic function $\mu(n)\Lambda(n)/\log(n)$ furnishes a characteristic function of primes. Importantly, it can be extended to localize onto prime k -tuples.

This characteristic function of primes allows for a direct representation of the exact prime counting function up to integer x

$$\pi(x) = - \sum_{n \leq x} \mu(n) \frac{\Lambda(n)}{\log(n)} .$$

Note this is not a Moebius inversion. The representation is readily extended to prime k -tuples. For example the exact twin-prime counting function is

$$\pi_2(x) = \sum_{n \leq x} \mu(n)\mu(n+2) \frac{\Lambda(n)\Lambda(n+2)}{\log(n)\log(n+2)} .$$

In addition to counting, the characteristic function suggests k -tuple analogs of the Chebyshev functions. Hopefully, possessing these explicit sums will enable further developments since they are constructed from well-studied objects. As a simple example we give a bound on the prime double Chebyshev functions averaged over prime doubles and show that Zhang's result implies they diverge as $x \rightarrow \infty$.

2 Counting k -tuples

Proposition 2.1 *Let \mathfrak{P}_k be the set of prime k -tuples, and denote a prime k -tuple by $\mathfrak{P}_k \ni \mathfrak{p}_k = (p, p + h_2, \dots, p + h_k)$ with $\mathcal{H}_k := \{0, h_2, \dots, h_k\}$ not necessarily admissible. The number of prime k -tuples up to some cut-off integer $x \geq p + h_k$ is given by¹*

$$\begin{aligned} \pi_{(k)}(x) &:= \sum_{\mathfrak{p}_k \in \mathfrak{P}_k}^x 1 \\ &= (-1)^k \sum_{n=2}^x \mu(n) \cdots \mu(n + h_k) \frac{\Lambda(n)}{\log(n)} \cdots \frac{\Lambda(n + h_k)}{\log(n + h_k)}. \end{aligned} \quad (2.1)$$

In particular, the number of prime doubles $(p, p + 2i)$ such that $x - 2 \geq 2i \in \mathbb{N}_+$ is

$$\pi_{(2)}(x) := \sum_{\mathfrak{p}_2 \in \mathfrak{P}_2}^x 1 = \sum_{n=2}^x \mu(n) \mu(n + 2i) \frac{\Lambda(n)}{\log(n)} \frac{\Lambda(n + 2i)}{\log(n + 2i)} \quad (2.2)$$

with twin primes corresponding to $i = 1$.

Proof: Since $\Lambda(n)$ restricts to prime powers p^ν while $\mu(p^\nu)$ allows only $\nu = 1$, then

$$\mu(n) \Lambda(n) = \begin{cases} -\log(p) & n = p \in \mathfrak{P}_1 \\ 0 & \text{otherwise} \end{cases}. \quad (2.3)$$

Loosely, $\mu(n) \Lambda(n) / \log(n)$ acts like a Dirac delta function for primes relative to the discrete measure on natural numbers. More precisely,

$$-\sum_{n=2}^x \mu(n) \frac{\Lambda(n)}{\log(n)} = \sum_{\mathfrak{p}_1 \in \mathfrak{P}_1}^x 1. \quad (2.4)$$

Simple induction on x proves (2.4) since it is obviously true for $x = 2$ and it jumps by one iff $x + 1 \in \mathfrak{P}_1$.

In general let $\mathbf{n}_k := (n, \dots, n + h_k)$, then

$$\mu(n) \Lambda(n) \cdots \mu(n + h_k) \Lambda(n + h_k) = \begin{cases} (-1)^k \log(p) \cdots \log(p + h_k), & \mathbf{n}_k = \mathfrak{p}_k \in \mathfrak{P}_k \\ 0 & \text{otherwise} \end{cases}. \quad (2.5)$$

Viewing \mathbf{n}_k as a point in a k -lattice and \mathfrak{P}_{k+1} as a subset of $\mathfrak{P}_k \times \mathbb{N}_+ = \bigotimes_k \mathfrak{P}_1 \times \mathbb{N}_+$,

¹The subscript (k) is supposed to indicate both the order k of the prime tuple and (implicitly) an associated $\mathcal{H}_k := \{0, h_2, \dots, h_k\}$. We will make the dependence on \mathcal{H}_k explicit when necessary.

the k -tuple result follows after observing that

$$\begin{aligned}
& \sum_{n=2}^x \left[\frac{\mu(n)\Lambda(n) \cdots \mu(n+h_k)\Lambda(n+h_k)}{\log(n) \cdots \log(n+h_k)} \right] \frac{\mu(n+h_{k+1})\Lambda(n+h_{k+1})}{\log(n+h_{k+1})} \\
&= (-1)^k \sum_{\substack{n' \leq x+h_{k+1} \\ \mathfrak{n}_k \in \mathfrak{P}_k}} \frac{\mu(n')\Lambda(n')}{\log(n')} \delta(n', (n+h_{k+1})) \\
&= (-1)^{k+1} \sum_{\mathfrak{p}_{k+1} \in \mathfrak{P}_{k+1}} 1 .
\end{aligned} \tag{2.6}$$

□

It is useful to introduce a more compact notation

$$\mu_{(k)}(n) := (-1)^k \mu(n) \cdots \mu(n+h_k) \tag{2.7}$$

and

$$\lambda_{(k)}(n) := \Lambda(n) \cdots \Lambda(n+h_k) / \log(n) \cdots \log(n+h_k) . \tag{2.8}$$

So we may write

$$\pi_{(k)}(x) = \sum_{n=2}^x \mu_{(k)}(n) \lambda_{(k)}(n) . \tag{2.9}$$

Now define the first and second Chebyshev functions for prime doubles;

Definition 2.1

$$\psi_{(2)}(x) := \frac{1}{2} \sum_{n=2}^x \lambda_{(2)}(n) \log(n(n+2i)) . \tag{2.10}$$

$$\theta_{(2)}(x) := \frac{1}{2} \sum_{n=2}^x \mu_{(2)}(n) \lambda_{(2)}(n) \log(n(n+2i)) . \tag{2.11}$$

There are obvious analogs of Chebyshev for higher k

Definition 2.2

$$\begin{aligned}
\psi_{(k)}(x) &:= \sum_{n=2}^x \lambda_{(k)}(n) \log(n_{(k)}) =: \sum_{n=2}^x \frac{\Lambda_{(k)}(n)}{\log^{k-1}(n_{(k)})} \\
\theta_{(k)}(x) &:= \sum_{n=2}^x \mu_{(k)}(n) \lambda_{(k)}(n) \log(n_{(k)}) = \sum_{n=2}^x \mu_{(k)}(n) \frac{\Lambda_{(k)}(n)}{\log^{k-1}(n_{(k)})}
\end{aligned} \tag{2.12}$$

where

$$n_{(k)} := (n(n+h_2) \cdots (n+h_k))^{1/k} \tag{2.13}$$

and

$$\Lambda_{(k)}(n) := \lambda_{(k)}(n) \log^k(n_{(k)}) . \tag{2.14}$$

Proposition 2.2

$$\theta_{(2)}(x) = \frac{1}{2} \sum_{p_2 \in \mathfrak{P}_2}^x \log(p(p+2i)) \quad (2.15)$$

Proof: Use the same reasoning as the previous proof. \square

We can obtain a tight bound on the average (with respect to i) prime-double Chebyshev functions. For example,

$$\begin{aligned} \widehat{\theta}_{(2)}(x) &:= \frac{\sum_{i=2}^{x-2} \theta_{(2)}(x)}{\sum_{i=2}^{x-2} 1} \\ &= \frac{1}{2} \frac{1}{(x/2 - 2)} \sum_{p_2 \in \mathfrak{P}_2}^x \left[\log(2^{x/2-2}) + \log(p^{x/2-2}) + \log\left(\frac{\Gamma(\frac{x+p}{2})}{\Gamma(\frac{4+n}{2})}\right) \right] \\ &\geq \frac{1}{2} \sum_{p_2 \in \mathfrak{P}_2}^x \left[\log(2) + \log(3) + \log\left(\frac{\Gamma(\frac{x+p}{2})}{\Gamma(\frac{4+p}{2})}\right)^{\frac{1}{x/2-2}} \right] \\ &> \frac{1}{2} \widehat{\pi}_{(2)}(x) + \frac{\widehat{\pi}_{(2)}(x)}{x} [\log(\Gamma(x/2 + 1)) - 1] \\ &= \widehat{\pi}_{(2)}(x) [O(\log(x)) + O(1)] . \end{aligned} \quad (2.16)$$

On the other hand,

$$\frac{1}{2} \sum_{p_2 \in \mathfrak{P}_2}^x \log(p(p+2i)) < \sum_{p_2 \in \mathfrak{P}_2}^x \log(p+2i) \leq \sum_{p_2 \in \mathfrak{P}_2}^x \log(x) = \log(x) \pi_{(2)}(x) . \quad (2.17)$$

So $\widehat{\theta}_{(2)}(x) \asymp \log(x) \widehat{\pi}_{(2)}(x)$. Because of Zhang's theorem [5], $\widehat{\pi}_{(2)}(x)$ must diverge with x . It follows that $\lim_{x \rightarrow \infty} \widehat{\theta}_{(2)}(x)/\log(x) = \infty$. Clearly the same bounds obtain for $\widehat{\psi}_{(2)}(x)$ in terms of $\widehat{J}_{(2)}(x)$ where $J_{(2)}$ is the weighted sum of prime-power doubles.

If the gamma hypothesis [6] holds, this can be strengthened:

Claim 2.1 *Assume the gamma distribution hypothesis and that \mathcal{H}_k is admissible. Then, for each $k \in \mathbb{N}_+$, $\pi_{(k)}(x) \sim c_{(k)} (x/\log^k(x))$ for some positive constant $c_{(k)}$.*

*Sketch of proof*²: As a direct consequence of the PNT;

$$\begin{aligned} \pi_{(1)}(x) &= - \sum_{n=2}^x \mu_{(1)}(n) \lambda_{(1)}(n) \sim x/\log(x) \\ &\Downarrow \\ \pi_{(1)}(N+1) - \pi_{(1)}(N) &\sim 1/\log(N) . \end{aligned} \quad (2.18)$$

²The idea of the proof is obvious and simple since the gamma distribution hypothesis eliminates the difficult part of the proof by fiat. Of course strict analytic rigor is necessary to promote this claim to a theorem: even then it would hold only conditionally.

Hence the density of log-primes goes like $\mu(n)\Lambda(n) \sim 1$.

Now, relative to the density of prime doubles, $\pi_{(2)}(x)$ represents a weighted intersection between the 2-lattice of integers and a ray $\mathbf{r}_{(2i)}$ along a direction dictated by $\mathcal{H}_2 = \{0, 2i\}$. Typically, $\mathbf{r}_{(2i)}$ will meet only coprime points due to the Von Mangoldt product so it is more efficient to restrict attention to the coprime 2-lattice. Explicitly, denote the coprime 2-lattice by the set of points $\{(n_1, n_2) \in \mathbb{N}_+^2 \mid \gcd(n_1, n_2) = 1\}$. Then

$$\begin{aligned} \pi_{(2i)}(x) &= \sum_{n=2}^x \mu_{(2i)}(n) \lambda_{(2i)}(n) \\ &= \sum_{n_1 \leq x} \sum_{n_2 \leq x+2i} \mu_{(1)}(n_1) \lambda_{(1)}(n_1) \mu_{(1)}(n_2) \lambda_{(1)}(n_2) \delta(n_2, n_1 + 2i) \\ &= \sum_{n_1 \leq x} \frac{\mu(n_1) \Lambda(n_1)}{\log(n_1)} \left(\sum_{n_2 \leq x+2i} \frac{c_{(2i)}(n_2|n_1)}{\log(n_2)} \delta(n_2, n_1 + 2i) \right). \end{aligned} \quad (2.19)$$

The delta function restricts to $\mathbf{r}_{(2i)}$ where n_1 and $n_1 + 2i$ are coprime. The weight function $c_{(2)}(n_2|n_1)$ encodes the condition $\mu(n_2)\Lambda(n_2) \neq 0$ given that $\mu(n_1)\Lambda(n_1) \neq 0$ along $\mathbf{r}_{(2i)}$.

Consider a change of summation variable $n_2 \mapsto n'_2 = n_1 + 2j$ in the inner sum with $2j = (2i)^l$ ensuring that n'_2 is also coprime to n_1 . This will generate a different ray $\mathbf{r}_{(2j)}$, but the intersections of $\mathbf{r}_{(2i)}$ and $\mathbf{r}_{(2j)}$ with the lattice are congruent in the sense that both intersect the lattice for the same set of n_1 because $\gcd(2i, 2j) = 2i$. In essence, $\mathbf{r}_{(2i)}$ gets shifted ‘up’ $2j - 2i$ units. But, according to the gamma hypothesis, the counting of prime powers is a random process. The two counting processes along $\mathbf{r}_{(2i)}$ and $\mathbf{r}_{(2j)}$ are independent, and they share congruent sets of coprime events and the same underlying distribution of prime powers. Consequently, for sufficiently large x , we expect $c_{(2i)}(n_2|n_1)$ to depend only on the equivalence class $[2i]$ defined by the relation $\gcd(2i, 2j) = 2i$. Hence, for sufficiently large N ,

$$\pi_{[2i]}(N+1) - \pi_{[2i]}(N) \sim \frac{1}{\log(N)} \frac{c_{[2i]}(N)}{\log(N)}. \quad (2.20)$$

Note that $c_{[2i]}(N)$ applies to counting numbers along $\mathbf{r}_{[2i]}$, and it detects prime doubles. It is decisive here that the assumed gamma distribution for prime doubles [7] allows this condition on counting numbers to be quantified as a probability statement implying that $c_{[2i]}(N) = O(1)$ for all i .³ In other words, every $\mathbf{r}_{(2)}$ has a non-vanishing intersection with the coprime 2-lattice for *any* sufficiently large region of the lattice, and the intersection density is asymptotically constant. So we end up with $\pi_{(2)}(x) \sim c_{(2)}(x/\log^2(x))$. The argument easily extends inductively, so $\pi_{(k)}(x) \sim c_{(k)}(x/\log^k(x))$ in general. \square

Aside from a few details, this sketch is essentially the familiar probabilistic argument so it doesn’t carry much weight. However, the exercise is useful because it is formulated in the context of a pair-wise coprime lattice structure which affords an unconventional perspective.

³Since we have assumed a probability model, the usual identification of $c_{[2i]}(N)$ with the singular series can be made.

References

- [1] G.H Hardy and J.E. Littlewood, Some problems of ‘Partitio numerorum’; III: On the expression of a number as a sum of primes, *Acta Math.*, **44**(1), (1923), 1–70.
- [2] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. Math.*, **167**, (2008), 481–547
- [3] D.A. Goldston, J. Pintz, and C.Y. Yildirim, Primes in tuples I, *Ann. Math.*, **170**(2), (2009), 819–862.
- [4] J. Korevaar, Prime pairs and the zeta function, *Jour. Approx. Theor.*, **158**, (2009), 69–96.
- [5] Yitang Zhang, Bounded gaps between primes, *Ann. Math.*, **179**(3), (2014), 1121–1174.
- [6] J. LaChapelle, Evidence of a Gamma Distribution for Prime Powers, arXiv:math.NT/1307.1717 (2013).
- [7] J. LaChapelle, Functional Integration on Constrained Spaces II: Applications, arXiv:math-ph/1405.0461 (2014).